

# AppaDB: an AcedB database for the nematode satellite organism *Pristionchus pacificus*

Jagan Srinivasan, Georg W. Otto<sup>1</sup>, Ulrich Kahlow<sup>2</sup>, Robert Geisler<sup>1</sup> and Ralf J. Sommer\*

Abteilung für Evolutionsbiologie, <sup>1</sup>Abteilung für Genetik and <sup>2</sup>Abteilung für Protein Evolution, Max-Planck Institut für Entwicklungsbiologie, Spemannstrasse 35–37, 72076 Tübingen, Germany

Received August 1, 2003; Revised and Accepted September 24, 2003

## ABSTRACT

*Pristionchus pacificus* is a free-living nematode of the Diplogastridae family and was recently developed as a satellite system in evolutionary developmental biology. AppaDB, a *P.pacificus* database, was created (<http://appadb.eb.tuebingen.mpg.de>) to integrate the genomic data of *P.pacificus*, comprising the physical map, genetic linkage map, EST and BAC end sequence and hybridization data. This developing database serves as a repository to search and find any information regarding physical contigs or genetic markers required for mapping of mutants. Additionally, it provides a platform for the *Caenorhabditis elegans* community to compare nematode genetic data in an evolutionary perspective.

## INTRODUCTION

*Pristionchus pacificus* is a free-living nematode of the Diplogastridae family and has been established as a satellite organism in evolutionary developmental biology (1–3). Satellite species are species that are sufficiently closely related to well-known model organisms that the genetic regulation of homologous cellular and developmental processes can be studied (1,3). *P.pacificus*, like the model organism *Caenorhabditis elegans*, can easily be cultured in the laboratory. It propagates as self-fertilizing hermaphrodites or by outcrossing after the spontaneous generation of males and has a 4-day life cycle (4). *P.pacificus* is amenable to various cellular, genetic and molecular techniques successfully used in *C.elegans*. Comparison of mutations in homeotic genes in both species revealed substantial changes in the function of these genes during development e.g. vulva formation (5–7). In recent years, developmental and molecular studies have been complemented by a genomic approach. Specifically, a genetic linkage map (8) and a physical map (9) were generated and two large-scale sequencing projects (EST and BAC end sequencing) performed. To compile all the genomic information generated in *P.pacificus*, we developed a database called AppaDB that presents all the information in a comprehensive manner. This database serves as a global resource to cover various aspects of genomics in *P.pacificus*.

## AppaDB DESIGN

AppaDB was implemented using ACedB (A *Caenorhabditis elegans* database) software designed by Richard Durbin (Sanger Centre, UK) and Jean Thierry-Mieg (CNRS, France). ACedB version 4.9 for Unix was used along with Perl 5.6 (10). The database was made accessible at <http://appadb.eb.tuebingen.mpg.de>.

ACedB was chosen as the software of choice for constructing this database because of the flexibility and usefulness in modelling biological objects as well as a user-friendly and graphical interface. AppaDB, like ACedB, is an object-oriented database, meaning the data are divided into several classes, which are linked to one another. The structure of AppaDB consists of (i) Clones, (ii) EST sequences, (iii) Sequence tagged sites (STSs, which includes SSCP markers), (iv) Probes (containing hybridization data), (v) Genetic Map, and (vi) Physical Map. Each of these classes is unique and contains several objects.

AppaDB has been designed to link different classes with one another. This integration allows the user to browse any information regarding the genetic linkage map or the physical map and presents all this information for query. The entire data set for AppaDB was initially in the form of ASCII flat files and was then converted to ACedB (ace) format files using Perl scripts. A survey of the various classes of AppaDB is provided below.

## CLONES

Two different BAC libraries (HindIII and EcoRI) of the *P.pacificus* genome were constructed (8). In total, ~19 000 clones from both the libraries were end sequenced and blasted. All this information regarding the BAC end sequences are compiled in the database. These BAC end sequences are complemented by an ongoing sequencing project of a fosmid library (R. Holt and R. J. Sommer, unpublished data).

## GENETIC LINKAGE MAP AND STSs

The genetic linkage map of *P.pacificus* was constructed using single-stranded conformational polymorphism (SSCP) markers generated from BAC ends and EST clones (8). *P.pacificus* has six chromosomes and the genetic linkage map consists of six linkage groups. All markers are incorporated into AppaDB using the marker name and chromosomal coordinates.

\*To whom correspondence should be addressed. Tel: +49 7071 601 371; Fax: +49 7071 601 498; Email: [ralf.sommer@tuebingen.mpg.de](mailto:ralf.sommer@tuebingen.mpg.de)

## PHYSICAL MAP

The physical map of *P.pacificus* consists of amplified restriction fragment length polymorphism (AFLP) fingerprints of >9000 BAC clones from both the libraries. The fingerprint information of each individual BAC was compiled in FPC v6.2 and contigs were made at fixed tolerance and cut-off values (9). Since the SSCP markers were mostly generated from BAC ends, the genetic linkage map information of these SSCP tagged BAC clones could be anchored directly on the physical map. The complete physical map was incorporated as an ace file into AppaDB.

## HYBRIDIZATION DATA (PROBES)

The BAC libraries were gridded on filters and these filters were hybridized with individual molecular probes. The BAC clones obtained were informative for positional cloning projects.

## EST SEQUENCES

An EST sequencing project was initiated in collaboration with the Parasitic Sequencing Project at the Washington University, St Louis (11). The entire EST data set can be downloaded from the NCBI. We have generated SSCP markers in several selected EST clones. These are represented in AppaDB.

## BROWSING AND QUERY INTERFACES

The AceBrowser interface in AppaDB allows complex queries to be addressed and displays the output information in a comprehensive way. AppaDB can be browsed in several ways:

(i) Browsing on classes: since the database is object oriented, one can browse through the various classes.

(ii) Searching on classes: the query form offers a convenient way to look for gene information, clone information, sequence information or hybridization data. Any information in a particular class is linked with information in other classes. For instance, if a SSCP marker has a corresponding contig, this information would be displayed and can be accessed by clicking on the contig.

(iii) Display: these data can be viewed either graphically or in a tree mode display.

## FUTURE DIRECTIONS

AppaDB was conceived to serve as a global resource to cover all aspects of *P.pacificus* genomics. Since the evolutionary lineage of *P.pacificus* diverged from that of *C.elegans* ~100–200 million years ago (3), it is a useful system to the *C.elegans* community. In particular, comparing the promoter regions of orthologous genes between *P.pacificus* and *C.elegans* can identify candidate regulatory control regions. Also in the years to come, the *Pristionchus* community is expected to grow with additional information being generated, such as whole-genome shotgun approaches. Therefore, AppaDB serves as a well-suited and flexible platform to provide these *P.pacificus* data to a broader audience.

## REFERENCES

1. Simpson, P. (2002) Evolution of development in closely related species of flies and worms. *Nature Rev. Genet.*, **3**, 907.
2. Sommer, R.J. (2000) Evolution of nematode development. *Curr. Opin. Genet. Dev.*, **10**, 443–448.
3. Eizinger, A., Jungblut, B. and Sommer, R.J. (1999) Evolutionary change in the functional specificity of genes. *Trends Genet.*, **15**, 197–202.
4. Sommer, R.J., Carta, L.K., Kim, S. and Sternberg, P.W. (1996) Morphological, genetic and molecular description of *Pristionchus pacificus* sp. n. (Nematoda: Neodiplogastridae). *Fund. Appl. Nematol.*, **19**, 511–521.
5. Sommer, R.J., Eizinger, A., Lee, K.Z., Jungblut, B., Bubeck, A. and Schlak, I. (1998) The *Pristionchus* HOX gene *Ppa-lin-39* inhibits programmed cell death to specify the vulva equivalence group and is not required during vulval induction. *Development*, **125**, 3865–3873.
6. Jungblut, B. and Sommer, R.J. (1998) The *Pristionchus pacificus* *mab-5* gene is involved in the regulation of ventral epidermal cell fates. *Curr. Biol.*, **8**, 775–778.
7. Eizinger, A. and Sommer, R.J. (1997) The homeotic gene *lin-39* and the evolution of nematode epidermal cell fates. *Science*, **278**, 452–455.
8. Srinivasan, J., Sinz, W., Lanz, C., Brand, A., Nandakumar, R., Raddatz, G., Witte, H., Keller, H., Kipping, I., Pires-DaSilva, A. et al. (2002) A bacterial artificial chromosome-based genetic linkage map of the nematode *Pristionchus pacificus*. *Genetics*, **162**, 129–134.
9. Srinivasan, J., Sinz, W., Jesse, T., Perebolte, L.W., Jansen, K., Buntjer, J., van der Meulen, M. and Sommer, R.J. (2003) An integrated genetic and physical map of the nematode *Pristionchus pacificus*. *Mol. Gen. Genet.*, **269**, 715–722.
10. Stein, L.D. and Thierry-Mieg, J. (1998) Scriptable access to the *Caenorhabditis elegans* genome sequence and other ACEDB databases. *Genome Res.*, **8**, 1308–1315.
11. McCarter, J.P., Clifton, S.W., Bird, D.M. and Waterston, R.H. (2002) Nematode gene sequences, update for June 2002. *J. Nematol.*, **34**, 71–74.